Measuring HCHO Oxidation Provides An Estimate of Ozone Production Rate for Urban and Regional Smog

Chatfield, R.B., NASA Ames Research Center, MS 245-5, Moffett Field, CA 94035 USA
Ren, X., RSMAS, Univ. of Miami
Brune, W., Pennsylvania State Univ.
Schwab, J., ASRC, SUNY Albany
Queens College and PMTACS Measurement Teams

Abstract:

Several indicators of smog ozone generation can be surprisingly well estimated using easily measured non-radical species like nitrogen oxides (NO_x), formaldehyde (HCHO), photolysis rates (UV), and temperature (T) using statistical relationships that exhibit some generality. We report an adaptation of an analysis developed for rural air to the atmosphere sampled in central Queens, NYC, with considerable freshly emitted pollutants. Direct measurements of the radicals [HO_2] and [OH] were available at this site, allowing direct estimation of $P_o(O_3) = k[HO_2][NO]$, $L(NO_2) = k[OH][NO_2]$, and the ratio of these rates (termed a production efficiency of O_3) in terms of non-radical species which are more easily measured. These provide approximations to the chemical ozone gross production rate, the loss of NO_2 by photochemistry, and an estimate of ozone production efficiency. In our analysis, R^2 for the estimated principal ozone O_3 production rate (involving HO_2) was in the range 0.48 to 0.74. Simple relationships appear to hold for high-NO very polluted cases. More complex relationships are found for intermediate polluted situations, where variations in volatile organic carbon (VOC) mixes or lack of an HCHO near-steady-state appear to complicate the relationships. T is a remarkably informative variable in this surface-level urban environment. Remaining variability in the statistical estimates of $P_o(O_3)$ and $L(NO_2)$ may be characterized in terms of gain factors, each varying slowly over periods of days. An understanding of these gain factors promises a more complete analysis of the $P_o(O_3)/ L(NO_2)$ and production efficiency. Additionally, the method again suggests a quantitative and very local application of descriptions of "VOC limitation" or "NO_x limitation" in photochemical smog relationships which allows us to explore for unexpected sources or photochemical behavior at field sites that need measure only HCHO, NO_x, and UV radiation. Complete meteorology-driven simulations are not replaced, but can be evaluated on their characterization of peroxy radicals. This may be valuable considering the complexity of sources, VOC reactions, and precise transport necessary capture peroxy concentrations by first-principles calculation.

Introduction

Understanding the origin of high levels of boundary layer smog ozone is a concern for health and agricultural economic reasons; concern has risen sufficiently to motivate a new, lower regulatory standard (US EPA1987, 2008). Lofting and long-distance transport of ozone and its precursors are also of concern with respect to ozone's role as a greenhouse gas and an intercontinental pollutant (Bey et al., 2002.)

The processes determining levels of smog ozone are complex, involving the chemical production, destruction, surface uptake, and transport of ozone, complicated all the more by similar processes determining the levels of its precursors. The non-linear

relationships change over time. The chemical production term for ozone is perhaps the most difficult to describe from first principles, since it depends on the production rate and levels of peroxy radicals, both HO_2 and organic peroxy radicals.

The ozone gross production rate is described by the dominant terms for the rate

$HO_2 + NO \rightarrow OH + NO_2$ (R1)

$P_o(O_3) = k_{HO_2+NO}[HO_2][NO]$

$$P(O_3) = k_{HO_2+NO}[HO_2][NO] + \sum_{i=1,n} k_{R\,HO_2+NO}[R_iO_2][NO]$$

where the number n of organic peroxy radicals can run to hundreds and thousands, some with very minor roles. The "principal" production rate $P_o(O_3)$ due to HO_2 radicals is very typically about 60% of the *total* production rate $P(O_3)$, according to calculations we have made with simple photochemical box models using a variety of condensed reaction mechanisms. Peroxy radicals can be measured only in specialized field campaigns, due to the technology required (Eisele et al., 1997, Brune et al. 1998). They are commonly estimated by calculation, but this involves many long, branching, and intricate reaction chains beginning from primary VOC's as they are initiated by attack by OH radical, O_3, or less commonly, photolysis or NO_3 attack. The number of reactions required to describe the concentrations of all secondary organic compounds runs potentially to tens of thousands of reactions (e.g. Jenkin et al. 1997. MCM v3.1, http://mcm.leeds.ac.uk/MCM/home.htt), with most kinetic rate processes determined by analogy with the relatively few reaction rates that have been measured with agreed-upon accuracy [Atkinson et al., 2005, Atkinson and Arey, 2003, Atkinson et al. 2006]. Beyond the uncertainties in the reactions and their rates frequently lie more uncertainties about quantitative VOC emission descriptions. As we noted, these interact with uncertainties in the transport of primary and secondary VOC's and reactive nitrogen oxides.

Almost as difficult to calculate completely as the ozone chemical production is the rate of another reaction involving HO_x and NO_x:

$OH + NO_2 \xrightarrow{M} HNO_3$ (predominant product) (R2)

which is given by

$R_{OH+N}O_2 = k_{OH+NO_2}[OH][NO_2]$

This term controls the removal of NO_x from the system, which in turn controls the reaction of peroxy radicals that allows further production of ozone.

In Chatfield et al (2009a), we suggested a diagnostic method for understanding smog. Using special-campaign information in which HO_2, NO, and accompanying species were measured, we found remarkably simple functional empirical relationships, and that functions of the simply measured variables were a good quantitiative guide to the ozone production rate due to HO_2. The variables were [NO] and the product j_{HCHO} x HCHO. A function of these two variables reproduced the chemical production of O_3, in a "contour plot" method reminiscent of the "empirical kinetic model approximation" method used in the 1980's (Dodge 1977, Shafer and Seinfeld, 1985), and of several portrayals since (e.g., Sillman and He, 2002). HCHO is our reasonable choice to trace organic oxidation and the production of HO_2 radicals because it is a near-terminal breakdown product in reaction chains involving a very large number of VOC emissions. Also, it is

short-lived enough to give a good local gauge of VOC oxidation activity over the previous few hours.

The method is best explained if we may assume that HCHO is not too far from a quasi-steady state slowly evolving with accompanying VOC's and also the photolytic radiation that produces it and help destroy it. We expected that urban emission would not have built up secondary HCHO, and that primary (or extremely rapidly produced) HCHO would be high enough that the technique would work in regions of large fresh urban emissions. Formaldehyde destruction rates vary strongly with UV radiation; indeed HCHO can persist at night away from surface. If one uses the product $j_{HCHO->HCHO+H}$ [HCHO],where the photolysis rate is the photolysis of formaldehyde to radicals, one introduces a timescale. This particular photorate is chosen since it is more energetic UV, and so it imitates the effects of the softer-UV photolysis to H_2 and CO, and also the formation of OH by harder UV processes, e.g., the formation of OH from $O(^1D)$ or the photolysis of higher aldehydes and peroxides. It was chosen to be intremediate. All photo rates tend to be closely linked, varying mostly in form due to ozone column and aerosol effects. Measurements of UV and can relatively easily produce estimates of this rate $j_{HCHO->HCHO+H}$ to useful accuracy. Remarkably, similar relationships involving j x HCHO and NO (as we shall abbreviate them below have shown some usefulness at high pollutant concentrations typical of New York City. These indications are appealing, since it does seem that j x HCHO should indicate VOC oxidation, but that the amount of HCHO should also depend somewhat on the VOC mix: i.e., toluene might produce one formaldehyde per 7 carbon atoms oxidized, isoprene might produce two to three for 5, ethane one or two for 2, etc. We almost always observe complex mixes of VOC's (Blake, INTEX-NA, VOC's at NYC), which should tend to produce some universality in the HCHO yield; still, some variation in the yield is to be expected. Kleinman (2004) points out that formaldehyde photolysis relates to two, partially distinguishable roles, as an indicator of fresh HO_x radicals and also as an indicator of general VOC oxidation, which also converts OH to HO_2 (and RO_2) radicals without creating new HO_x.

Our approach to an empirical description concentrates on a local description rather than the time-integrated ozone production from a strong source. Two previous approaches describing integrated production are particularly userful as they develop themes that are also useful here. One approach relied more on kinetic models applied in idealized situations (a single typical VOC-NO_x mix characterizing a city or region) (Kleinman, 2004; Kleinman et al., 2000). Another approach relied more on by-products of ozone production (e.g., H_2O_2 and HNO_3), and has been used to characterize the fate of HO_2 radicals, based on correlations of HO_2 and OH radical reactions directly involved with ozone production and those that produce these radicals. (Sillman and He, 2002).

For this sampling site the estimation works well in the broadest view, but can be improved for more detailed analysis. The goodness of fit that we find using our procedure proves to hide some complications that are revealed when we restrict our consideration to situations with polltuant levels of NO less than 8 ppb. We will concentrate on the more difficult complex dataset, and find succeeding levels of approximation to the principal ozone production rate. Certain consistencies persist when the best fits vary systematically

based on just one parameter that varies over the period of one or two days independent of the other pollutant measures. We seek some explanation for this variation. Similar slow variability is found in fits made to the NO_2 disappearance rate R_{OH+NO2} . There are separate parameteric variations controlling R_{OH+NO2}, and these seem to follow some measures of VOC reactivity revealed by measurments of hydrocarbon species.

Sampling Location and Instrumentation

Sampling was conducted from a small scaffolding tower located at Queen's College, which centrally located within Flushing, New York City; the PM2.5 Technology Assessment and Characterization Study (PMTACS) study concentrated on aerosols deriving from the city and regionally upwind. The gas-phase species concentrations were characteristic of the strong local sources, but were lower when east winds brought air from the Atlantic. The site has commercial, residential and industrial sources, and is ringed by heavily trafficked freeways and arterials. Relatively high levels of NO_x and anthropogenic VOC's characterized the area. Further details given in the Ren et al. (2003a) publication describing radical measurements.

Ren et al. (2003a) detail the laser-induced fluorescence measurements of both HO_2 and OH that were used to quantify ozone production using the Pennsylvania State GTHOS (ground-based radical-measurement instrument). NO, NO_2, CO, O_3, SO_2, CH_4 and other species were measured by the New York City using one of their PAM (Photochemical Assessment Monitoring Station) sites (US Environmental Protection Agency — EPA, 1994). Formaldehyde was scrubbed from ambient air into water and then react to produce 3,5-diacetyl-1,dihydrolutidine under the care of co-author James Schwab. Other carbonyls were also measured every third day with three-hour integrated sampling. VOC's were measured every hour, but at a distance 2.5 km away at Queensborough Community College. Histograms of pertinent variables to this paper are shown in Figure 1(a). Figure 1(b) decribes an interesting subset defined and used in a later section on "moderate pollution." Ren et al (2003b) describe techniques to check the appropriateness of the VOC measurements and also the estimation of photolytic dissociation rates based on a multiwavelength shadowband UV radiometer. That paper makes comparisons of measured and model-estimated HO_2 and OH. Some considerations of simulated and observed radical concentrations should be borne in mind while considering our work, although, from a broader perspective, the implications are still not completely clear. The simulations followed HO_2 and OH concentrations well over most hours of the days July 10 to August 2, 2004. Ren et al. (2004) note however systematic differences between modeled and measured radical ratios.

Modeled HO_2 /OH ratios exceeded measured HO_2 /OH ratios substantially in conditions when NO was below 1 ppbv, but exceeded them above when NO was 8 ppbv. Ren et al. (2003b) summarize the situation: "This difference is consistent with measured OH being greater than modeled OH at low NO, while measured HO_2 is much greater at high NO." Other papers by these authors find similar but not identical variations in other, cleaner-air sampling conditions.

Organic Reactivity, NO, and Empirical Estimates of $P_o(O_3)$

Following ideas from work on regional smog ozone (Chatfield et al., 2009), we sought relationships between [NO], our measure of organic reactivity j **x** HCHO and $P_o(O_3)$. Unless otherwise noted, the photolysis rate used j is always the photolysis rate of HCHO to radicals, and responds to hard-UV radiation, but not as hard as for $O(^1D)$ from ozone. Therewe found adequate a very simple relationship *constant* x (j HCHO **x** NO])^a, where a ~ 0.4, and the product j HCHO **x** NO was below ~0.4 ppb^{2}hr^{r1} and the production of O_3 from HO_2 was below ~10 ppb hr^{r1}. It seemed reasonable that organic reactivity and nitrogen oxides might influence $P_o(O_3)$ differently at higher concentrations, so we sought a more general relation:

$$P_o(O_3) = C \ (j \ \mathbf{x} \ [HCHO]^a \ [NO])^b$$

or,

$$\log P_o(O_3) = \log C + a \log (j \ \mathbf{x} \ [HCHO]) + b \log [NO]$$

The latter form allows a simple linear regression, and is a good form, since residuals are most nearly normally distributed.

This formulation has proved reasonably successful when considering both polluted and very polluted situations, and should allow a simple asymptotic understanding in any case. If j **x** HCHO is considered to reflect the oxidation rate of volatile organics, the formula should be most appropriate when HCHO is in an evolving steady state with organic precursors and its OH and photolytic sinks; this is conveniently estimated using the sink terms, $t = 1 / (k_{OH+HCHO}[OH] + j_{HCHO->rads} + j_{HCHO->H2})$ where the j's are appropriate photolysis rates and $k_{OH+HCHO}$ is the rate coefficient for reaction with hydroxyl. All these sink terms are at maximum near midday. To allow for a close, not greatly time-delayed, relationship between radicals and $P_o(O_3)$, we made empirical statistical models of the data obseved between 10 AM and 4 PM local daylight time (approximately 9 AM and 3 PM sun time). Since pollution levels, especially [NO], were occasionally extremely high, it seemed appropriate to consider two situations, "All Pollution," and "Moderate Pollution." Recognizing the variability of pollution situations (NO amounts, VOC amounts and speciation), we separated the "Moderate Pollution" case with the hypothesis that it might serve as a guide to other moderately polluted situations. "All Pollution" cases were hypothesized to be more appropriate for the Queens, NYC sampling situation, but perhaps a useful glimpse at fresh, less-diluted situation.

Results and Discussion: $P_o(O_3)$ Including Heavily Polluted Situations.

We begin with an example in which the empirical technique appears to explain the observed variance very well. Somewhat counter-intuitively, highest correlations of observed and modeled $P_o(O_3)$, occur when we parameterize the whole dataset rather than examine subsets. This urban dataset has a large variance and the contrast between very high ozone production rates and moderately high rates dominates the fitting procedure and the correlation of observations and empirical fit. The first test included all pollution samples at the Queens College site during photochemically active hours (10 AM to 3 PM). The NO mixing ratio reached a maximum of 25 ppb! Clearly very recent pollution was

being sampled, and it is doubtful that the time from formaldehyde emissions to sampling was large compared to the photochemcial equilibration time (lifetime) of the species. Nevertheless, the estimated empircal fit was good, with

All pollution:

$a = 0.419 \pm 0.006$ for HCHO activity

$b = 0.548 \pm 0.006$ for NO

$c = -0.42 \pm 0.07$

Mulitple R^2 (for log-concentration quantities) = 0.73

R^2 (for concentration-unit quantities) = 0.71 based on 4911 samples.

All estimates were formally determined to a *p*-value of $< 2 \times 10^{16}$. Contour-plots of the dependence of $P_o(O_3)$ on the defining chemcal variables is shown in Figure 2. These estimates for *a* and *b* are relatively close to estimates made for airborne sampling of urban, rural, and some remote boundary layers in separate Central-US/East-Coast and California sampling (Chatfield, et al., 2009). The figure shows three depictions of the relationship to *j* **x** HCHO and NO, a two-d surface, a contour plot showing the same relationship, and a graph drawn when all three quantities are expressed in terms of $\log_{10}$(quantity).

The use of log values for predictors and predicted O_3 greatly simplifies the estimation of the relationships as well as their display. It is advisable to use log values in the regressions, since the log-based residuals to the regressions we perform behave approximately normally, while the non-log quantities yield residuals that are proportional to the predictors and predicted values. This may indicate that the "measured" Po(O_3), or perhaps the predictor variables, have errors that are proportional to their value. If this is indeed the case, then there is a limit to the expected goodness of fit we should expect from any statistical fits, a limit due to large variance of $P_o(O_3)$ (or contributing variables) that is beyond what may be expected from statistiics. The relationships are also expressible as sensitivities: $\partial \log P_o(O_3) / \partial \log (j \times \text{HCHO})$ and $\partial \log P_o(O_3) / \partial \log \text{NO}$.

The slope of the isopleths in the log-log plots also summarizes the ratio of the sensitivity to NO and HCHO activity. Diagonal 1:–1 lines on the log-log plots in this report indicate equal exponents or equal sensitivities for the control variable selected. Spacing of the contours is an indicator of the power law followed.

Comparing the linear and log depiction of $P_o(O_3)$ the log depiction is more intuitive and easy to read. However, the prominent feature of a central mound indicating "most efficient mixtures for ozone production" seems to disappear. This is only appearance. While equal horizontal increments in figure 4(a) can be interpreted as equal additions or subtractions of ppb, equal increments in figure 4(b) are equal *relative* changes in concentration (i.e., 50% reduction). (In this discussion, it is assumed that physical conditions for the UV-dependent *j* will remain the same for changes in organics that lead to increments in HCHO.) In other words, the "optimum corner region" for strong increase or decrease in ozone production should depend on the relative costs of reducing either VOC's or NO.

The use of the power law formulation is not derivable from the mathematics of the chemical kinetic system, but it has some advantages. The formulation describes with one parameter per control variable the common observations about smog ozone, which is that

instantaneous ozone *production* responds positively to increased precursor, but does so at an ever decreasing rate. Most other simple functional forms do not have these properties. (We ignore NO "titration" of ozone since it simplifies the discussion, and an "oxidant" quantity, $O_3 + NO_2 + O(^3P) + O(^1D) + 2\ NO_3$, is conserved. Oxidant is instead considered to be destroyed by reactions like $OH+NO_2$ or O_3+HO_2 or $O(^1D)+H_2O$) Another advantage of the formulation is the simplicity of the log-log graph system that results, which directly expresses sensitivities. Usually we found that more complex formulations did not add significance with added parameters. However, the use of a simple log formulation leads to somewhat different values of the exact parameters for observation datasets that contain very high pollution, high pollution, and regional dilute pollution. (Parameter estimates did vary in important subsets of this in the Queens dataset, as seen below. The relationship of these must be left to future work.

Once these valuations are set, a log or a non-log diagram can be used to make a simple estimate of an optimal control strategy, at least based on these local observations. True determinations of the control strategy of course depend on the effect of upwind nitrogen oxide emissions on NO, of VOC emission mixtures on HCHO. Indeed they depend on the integral production of ozone over all areas influencing observed concentration at a particular point and time. We will return to interpretation later.

$P_o(O_3)$ in More Complex, Less Heavily Polluted Situations

When we look at a significant subset of the same Queens dataset, the goodness of fit decreases and the relationships parameterizing ozone production become more complex. For example, we restricted NO to lower concentration (≤ 7 ppb). Since the emissions have had some reaction time, there should be greater likelihood that HCHO has come closer a production/loss quasi-steady state, and also there has been more mixing of multiple VOC types to a more "typical urban mix"; both considerations would suggest that statistical fits would be more accurate. Instead, there appear to be multiple similar relationships that suggest a remaining, elusive parameter. Figure 3 shows a series of correlation plots for successively more complex statistical relationships that are modifications of the basic power-law formula we have used so far. The first panel describes a correlation plot based only the formaldehyde activity and NO. Succeeding graphs describe possible estimation improvements which are introduced in later sections.

We summarize our findings in Table 1. This table allows for relationships with previously described VOC activity and NO, but adds temperature, and also a "specific situation" which is characterized a slow (several-day) variability with time.

Table 1 shows that, at first glance, ozone production is only modestly well predicted by our VOC-activity and NO variables.

$$P_o(O_3) = c\ \ (j \times [HCHO])^a\ [NO])^b$$

$R^2 = 0.41$ is indicated for a relationship based on VOC activity and NO alone, as used above. Despite this low correlation coefficient, there is remarkable uniformity on individual days and even on periods of a few days, as the first panel of Figure 3 shows.

Apparently ozone production relationships vary day by day, but have similar linear relationships with different slopes from day to day. Differing *slopes* of the non-log quantitities tend to follow different near-horizontal "layers" corresponding to the addition of different values of log(*slope*) under the conditions of Figure 3.

High temperatures have long been associated with high ozone concentrations and by extension high ozone production (Leighton,1961). Chamber studies suggest that this is more important in cases of higher NO_x additions compared to VOC additions {Carter et al., 1997). The Queens study fits this description. We considered the relationship

$$P_o(O_3) = c\ \ (j \times [HCHO])^a\ [NO])^b 10^{d(T\,/\,298)}$$

The temperature effect $10^{d(T/298)}$ was chosen as the numerically most tractable: Arrhenius formulations $\exp(-T_{Arrh}/T)$ and power formulations like T^d yielded products of huge and tiny numbers, while there was no gain in explanatory power.

Figure 3 (second panel) indicates how much correlation graphs tend to narrow when temperature is added. However, the upper end of the log plot splays wider than in the first panel, showing rather poor fitting at high values of $P_o(O_3)$ that has a large effect in limiting the explained variance of the physical (non-log) quantities to about 0.57.

Figure 4 (a) shows the relationship which includes T/298. Figure 3 (third panel) indicates how much correlation graphs tend to narrow when temperature is added. However, the upper end of the log plot splays wider than in the first panel, showing rather poor fitting at high values of $P_o(O_3)$ that has a large effect in limiting the explained variance of the physical (non-log) quantities to about 0.57. Analyses of the residuals for the first relationship (without *T*) suggested that a higher exponent, $(j \times [HCHO])^{\mathbf{0.42}}$ to $(j \times [HCHO])^{\mathbf{0.47}}$ would fit better at higher concentrations.

Let us consider why including temperature helped. It is important to remember how regression studies act when explanatory variables are somewhat correlated. Importance in the regression may appear to shift from one variable to a new variable when the latter is added. For example in Table 1 below, when temperature is added, the coefficient *b* for NO rises from 0.41to 0.58, while *a* for HCHO activity drops from 0.35 to 0.18. Figure 3 (e) shows the clear but complex relationship between temperature and the logarithm of HCHO activity. The correlation of *j* with *j* x HCHO is about 0.5. The correlation coefficient between *T* and (to *j* x HCHO) is ~0.4, but the graph implies a greater degree of relationship. Dropping data from just one day near the beginning of measurements would likely yield regressions with (*j* x [HCHO]) exponents more like the all-data case and rural measurements.

One reasonable explanation is that days with different temperatures may have other variations in chemical composition. Another possible reason is that *T* may more accurately reflect HCHO effects at the points that determine the recycling radicals HO_2 and OH measured at the trailer. Even more likely, very fresh emissions contain higher aldehyde products, or fresh alkenes or similar compounds that have yet to be processed to produce the high quantities of HCHO that they would have in a few-hours steady state. If other sources of HO_x and so HO_2 dominate the composition in certain cases, and temperature is linked to these variables temperature may dominate reactivity, and (*j* x HCHO) may help describe the activity of species emitted earlier.

In any case, the estimated effect of T is quite strong, as Figure 5(a) indicates. The combined effect of j x HCHO and T is close to the effect of NO over the range of values observed at Queens College.

There is some information from another sampling expedition for which we have made a preliminary analysis. Sampling during the California-ARCTAS intensive study included many airborne passes over Los Angeles or other South Coast cities. (ARCTAS: Arctic Research of the Composition of the Troposphere from Aircraft and Satellites; the California investigations were added into this airborne intensive after the name was fixed.) We used this Southern California geographical location and a condition that NO_x be above 2 ppb to define a somewhat analogous urban mix. Approximately six samples out of 150 showed a dramatically different pattern from the rest and were excluded. In this sample, the fitted relationship was

$$P_o(O_3) = 10^{-.51} (j \times [HCHO])^{0.69} [NO]^{0.29}$$

This relationship fitted with an R^2 of 0.69 in log quantities and 0.63 in non-log quantities. Temperature did not add any explanative power. (The NO and HCHO were measured by respective Weinheimer and Fried groups, and photolysis by the Shetter group (Singh et al., 2009)). While the airborne samples deserve more detailed study, the exponent for the VOC reactivity, is greater than several other estimates, 0.36 for the corresponding unelaborated response for Queens moderate pollution, 0.47 for moderately high pollution, and 0.42 for all Queens samples. Figure 5(b) shows a non-parametric local Krig-ing fit (*R* routine *krig.image* in the *fields* package). The 0.4 power law looks generally appropriate, although there is lots of variation, presumably reflecting the fresh emissions. By comparison, rural North American analyses gave an exponent of 0.4. We find these power-law estimates remarkably similar given the different instruments and sampling situation.

A note on the confidence intervals described. Essentially, selected subsets omitting cases are used to examine the coherence of a day-by-day progression in the linear relationship in a method called generalized cross-validation (Wood, 2007).

Day-by-Day Non-Predictive Analyses of $P_o(O_3)$

Figure 3 shows a comparison of observed and predicted ozone production as a function of sampling day. The striking thing to note is that different near-linear relationships are observable for different days. This led us to a final statistical fit, one which takes us beyond an obvious predictive capability. Based on our observations of different slopes on different days, we allow a slow progression with time, with the variability limited by the analysis method of *mgcv*.

$$P_o(O_3) = c \ (j \times [HCHO])^a \ [NO])^b 10^{d(T/298)} \ \boldsymbol{f}_{spline}(t) \ , \quad \text{i.e.,}$$

$$P_o(O_3) = 10^{-0.54} \ (j \times [HCHO])^{0.28} \ [NO])^{0.60} \ \boldsymbol{f}_{spline}(t)$$

Table 1 $P_o(O_3)$ for Moderately Severe Pollution Precursor Conditions

C	10^c	**1.16** ± 0.14	**–16.2** ± 0.12	**–0.54** ± 0.12	**–13.1** ± 0.63

a	(to j x HCHO)[a]	**0.36** ± 0.01	**0.17** ± 0.01	**0.28** ± 0.01	**0.17** ± 0.01
b	**(NO)**^{b}	**0.41** ± 0.01	**0.58** ± 0.01	**0.60** ± 0.01	**0.64** ± 0.01
d	$\mathbf{10^{d(T/298)}}$		**16.51** ± 0.38		**12.82 ± 0.64**
$f(t)$	**Day-dependent** $\mathbf{10^{f(t)}}$			**9.0** eq. dof	**9.0** eq. dof
	$\mathbf{R^2}$ **in log term**	0.65	0.75	0.75	0.79
	In ppb terms	**0.46**	**0.56**	**0.64**	**0.67**

The fitting parameters for the regression fit shown are shown in Table 1, yielding

390 *i.e.*, $P_o(O_3) = 10^{-13.1}$ $(j \times [HCHO])^{0.17}$ $[NO])^{0.64}$ $10^{12.82\,(T/298)}$ $\boldsymbol{f}_{spline}(t)$,

where t is measured in days and $f_{spline}(t)$ is a spline function estimated to minimize the residuals but simple enough to satisfy the generalized cross-validation condition. An additive model in the package *mgcv* was used (Wood, 2007). Figure 3 shows a much better fit that we obtain when we use day number as a regression variable, in a method that recognizes this slow variation. To what extent does this represent "over-fitting"? The function described has three maxima and the statistical fit routine describe it as having about 9 degrees of freedom (eq. dof) over 23 days. This number of degrees of freedom is still much less than the number of days of observations and extremely small compared to the number of observations (4261 observations made over). Sequential concentration

400 observations during the day show much more variability than this. Even accounting for some correlation of the observations in time, the information from the data greatly outweighs the number of parameters to be fit. We have examined the variation, and it has no clear relation to wind direction or classifications of hydrocarbon compounds (e.g., very reactive organics), and the correlation with the (T/298) is low. Unfortunately, there is *currently* no way of describing the spline function for future observations, so predictivity has been lost for this particular parameterization. We include the analysis here so as to stimulate future investigations that may explain this uniformity and so to understand the underlying chemistry better.

Several attempts were made to explain the slow variation $\boldsymbol{f}_{spline}(t)$. They are

410 somewhat instructive in what hypotheses they discourage. It was natural to seek relationships with hydrocarbon reactivity and tendency to produce HO_2 or RO_2 radicals. We sought this with a partial least squares (L1 or "lasso" variety) regression fit. Groups of variables, components, were sought that were most strongly related to the function $\boldsymbol{f}_{spline}(t)$. (The procedure resembles principle components analysis, except that orthogonal groups are found and ranked according to their contribution in explaining $\boldsymbol{f}_{spline}(t)$. This method has good properties in bridging between complete fitting and a very conservative "no fit" (all contributions zero) outlook, and it allows for cross-validation of estimates by selective data denial (Hastie et al., 2001). Figure 6 (a) shows the pattern that could be fit on a dataset

which included hydrocarbon concentrations. There was significant contribution of the first component, but little understanding: the compound identified as *trans*-2-butene was the major positive contributer (more ozone production), while *cis*-2-butene was distinguishably negative; 3-methyl pentane was the strongest negative contributor. Few components contributed to this group. This suggests, if anything, that particular sources had a strong effect, but not necessarily via the chemistry of these compounds. In view of this, a regression against wind direction was made. Two small maxima and minima were suggested, but the explanatory power was low. These investigations do not seem to tell us much, but we will return to the technique in the next section, where we discuss Figure 6 (b).

The OH+NO_2 Loss Term for Oxidant

There is another feature that also suggests further directions for study. Earlier we noted that most ozone destruction terms are relatively easily estimated as checks to models, but one term that typically does require photochemical modeling is the OH+NO_2 term. This is more properly a loss term for "odd oxygen" defined as the sum of rapidly interechanging species like O_3 and NO_2 [Kleinman, 2004, Seinfeld and Pandis, 2004]. To the extent that this loss term for odd oxygen is dominant photochemically, it has also been used to describe an "ozone production efficiency." This has been described as approximately the amount of ozone produced by an emitted NO molecule divided by the NO_x loss rate: how much odd oxygen is produced in a "cycle" of NO_x emission and removal.

Consequently, we wrote a relationship involving NO_2 similar to the ozone production term.

$$R_{OH+NO_2} = k_{OH+NO_2} [OH][NO_2]$$

We tried very similar fitting procedures; we used [NO_2] rather than [NO_2] since the latter can either be measured directly or estimated from the more easily measured [NO] and [O_3] by the famous Leighton [1972] relation. Using logarithms once again to make estimates (in the most detailed form)

$$\log R_{NO2+OH} = \log C + a \log (j \times [HCHO]) + b \log [NO_2] + d \log (T / 298) + \boldsymbol{f}(t) ,$$

We derived estimates (compared in Table 2 accompanied by approximate error estimates) such as the following

$$R_{OH+NO2} = 10^{0.42} (j \times [HCHO])^{0.17} [NO_2]^{0.90} ,$$

$$R_{OH+NO2} = 10^{1.51} (j \times [HCHO])^{0.11} [NO_2)^{0.70} 10^{0.048(T / 298)},$$

$$R_{OH+NO2} = 10^{1.61} (j \times [HCHO])^{0.21} [NO_2])^{0.53} \boldsymbol{f}_{spline}(t) ,$$

$$R_{OH+NO2} = 10^{1.93} (j \times [HCHO])^{0.16} [NO_2])^{0.52} 10^{0.018 (T / 298)} \boldsymbol{f}_{spline}(t)$$

The estimated coefficients follow generally similar patterns to those for $P_o(O_3)$. In this urban situation, measured HCHO activity and NO2 provide modest information about the variance of $L(NO_2)$, $R^2 = 0.47$. There is a notable increase in explanation when using T, but fitting a day-to-multi-day function adds considerably more to the R^2 value. Explanatory power (regression coeeficient magnitude) tends to redistribute from one independent variable to another as more variables are added. The increase in one regression coefficient for a new variable at the expense of regression coefficient of previous variable regressiion

coefficient may or may not necessarily imply conclusions about causation, of course. They do indicate specific ways to study the importance (for example) of temperature or formaldehyde activity by other techniques. Figure 7 (a and b) give an overview of sensitivities to radical production and NO_2 concentration in practical ppb units and also with a more abstract scale in which log quantities are used. The main contours indicate response to radical activity and to NO_2, while the lighter contours indicate comparison to the ozone production relationships just shown. This relationship is actually rather complex, depending on a strong simplifying assumption that NO/NO_2 is at the mean value observed in the samples, $NO/NO_2 \sim 0.21$.

Again, very significant increases in model R^2 were found when we allowed a separate, slowly varying, function of time to modify the other effects in the prediction in a multiplicative manner. R^2 of 0.88 was achieved. Figure 6 (b) describes this function; since the dimensionless function varies from by a range of 0.6 units, the sensitivity effects were larger than for $P_o(O_3)$. In this case, the *L1* partial least squares analysis suggested that there were strong negative effects on $L(NO_2)$ that were associated with high reactivity VOC compounds. Possibly OH was being reduced by the availability of other reactants (e.g., VOCs) than NO_2. If so, there should be ways to parameterize the $L(NO_2)$ response in a predictive manner. Measurement and kinetic calculations with observed VOC's couldbe an aid in this area near sources.

Table 2 $OH+NO_2$ Rate for Moderately Severe Pollution Precursor Conditions

c	10^c	**0.42** ± 0.10	**-21.8** ± 0.47	**1.61** ± 0.06	**–4.5** ± 0.49
a	$(j \times HCHO)^a$	**0.17** ± 0.02	**0.11** ± 0.01	**0.21** ± 0.01	**0.16** ± 0.01
b	$(NO_2)^b$	**0.90** ± 0.03	**0.69** ± 0.02	**0.53** ± 0.01	**0.52** 0.01
d	$10^{d(T/298)}$		**20.37** ± 048		**4.78** ± .51
f(t)	**Day-dependent** $10^{f(t)}$			**9.0** eq. dof	**9.0** eq. dof
	R^2 **in log term**	0.47	0.68	0.91	0.91
	in ppb terms	0.41	0.56	0.89	0.88

Approximate Efficiency of Ozone Production

Another significant estimator is the efficiency of ozone production. This is generally described as the total chemical production rate of ozone ratioed to the loss rate of NO_x from the system. An approximate simple form for this production efficiency can be described as the ratio of two rates, that of peroxy radicals reacting with NO to generate new ozone, $P(O_3)$, and the rate of loss of NO_2 by reaction with OH, i.e., $P_o(O_3) / L(NO_2)$. The values derived from the rate laws vary between 0.5 and 18, with one outlier at ~45. Statisticl Estimates of $P_o(O_3) / L(NO_2)$ without extra information can be made (See Figure

8(a)) but they describe 51% of the variance or less, although suggesting further relationships. One might expect that the NO/NO_2 ratio is very important for this ratio, and that a smaller role would be played by mechanisms determining ra8dicals, since OH and HO_2 would rise almost proportionally at the same NO/NO_2 ratio. Indeed, this is the case. Radicals play a certain role, though. The best estimate we found for this ratio in using familiar variables included only photolytic UV (the photolysis of HCHO to radicals), not HCHO.

$$\text{Ozone production Efficiency} = 10^{1.830} j^{1.830}\, NO^{0.81} NO_2^{-0.95},\ R^2 = 0.51$$

$$\text{or } 10^{-8.2} (j^{-0.2} NO^{0.82} NO_2^{-0.94})^{b} 10^{9.6(T/298)}\, \boldsymbol{f}_{\text{spline}}(t)\ ,\ R^2 = 0.78$$

The function $\boldsymbol{f}_{\text{spline}}(t)$ used to improve this estimate is nearly the opposite of the function describing NO_2 loss, as Figure 8(c) shows. Since the removal rate is so significant, it is quite possible that improvements in production efficiency parameterization can be made . This could occure if measurement of useful underlying variables such as VOC composition or VOC reactivity with OH do prove to be useful in descring $L(NO_2)$

Discussion and Comparison to Previous Work

There are many similarities of these relationships to those described by Kleinman (2004). Kleinman describes relationships for ozone produced along a parcel trajectory with power laws that look somewhat similar to our own. In these relationships, Kleinman makes a useful distinction between "original" sources of HO_x radicals and a conversion of OH to HO_2; he also includes a term for the conversion of NO to NO_2. We have not distinguished between the first two terms; Kleinman also notes that there is some overlap in the concepts, particularly the role played by HCHO. HCHO is both an important source of HO_x radicals and also a signficant measure of the rate of organic reactivity which converts OH to HO_2. Kleinman's (2002) suggested measurement-based description of ozone production efficiency points out the importance of [NO] / [NO_2]For that reason, the two descriptions must differ: perhaps Kleinman's description forces the distinction a bit strongly, whereas this description i bound to ignore what distinction there is. Our research is somewhat difficult to compare in detail to that of Sillman (Sillman and He, 2002, and referenced previous work) due to the near-source sampling. Sillman's analysis moved forward the longitudinal studies of ozone, nitrogen oxides, and some pollution tracers in the work that followed the Liu et al (1987) studies in NOAA Boulder's air chemistry group. Ren's point-modeling work on the Queens situation is closer in spirit, and the Sillman and the Ren papers provide a good companion to this. Possibly the greatest use can be in future comparisons to regional model simulations, particularly in their treatment of peroxy radicals and their behavior at high NO_x concentrations. This type of analysis may be greatly improved by VOC reactivity measurements wherever such equipment can be deployed (Mao et al, 2009).

Conclusions

We have found that combined measurements of formaldehyde, formaldehyde photolysis rate, and nitrogen oxides (NO and preferably NO_2) allow useful statistical estimates of the ozone production rate $P_O(O_3)$ and the loss of nitrogen oxide $L(NO_2)$ from a

photochemically reactive system. Specifically, the principal rates HO_2 +NO and OH+NO_2 have relationships to these relatively easily measured parameters, a radical activity parameter involving VOC, (to j x HCHO), and the appropriate nitrogen oxide concentration, NO or NO_2. The relationships are always significantly less than linear with the NO_x quantity, so that the regression studies add information. As with most other considerations of pollutant ozone formation, $P_o(O_3)$ always increases with added precursor, but at an ever decreasing rate. (Concentrations of NO_x at power-plant plume levels may not seem to follow this pattern; that topic deserves separate discussion.) A power-law form with the power a or b less than 1 provides appropriately simple fitting parameters. Log-log graphs with the power form give a useful tool, e.g., simple, linear contour graphs which illustrate sensitivities. When looking at the largest span of polluted situations including very fresh pollution with 8-50 ppb of NO, the variance explained for $P_o(O_3)$ is remarkable, reaching 0.73. When more modest NO values between about 0.5 and 8 ppb are examined, the scatter in prediction remains, while ozone varies over a shorter range, lowering the correlation coefficient. For the moderate-pollution example, temperature seems to very play an important explanatory role, a situation not found with airborne studies we consider more fully elsewhere (ref). For $P_o(O_3)$, nitric oxide concentrations and another variable, VOC activity or temperature, play roughly equal roles over the range of variables, NO, to j x HCHO , and T encountered in the sampling situation. For the loss of oxidant due to the reaction of OH + NO_2 -> HNO_3, similar relationships are found. For each relationship, $P_o(O_3)$ and $L(NO_2)$, there seem to be relationships that change the sensitivity to the other variable over periods of several days. For $L(NO_2)$, the relationship appears related to the presence of highly reactive hydrocarbons, possibly indicating a competition for OH radicals. The origin of the day-to-day variation in $P_o(O_3)$ is not yet clear, but it seems to be a simple parameter, a general amplification or diminution of the the $P_o(O_3)$ response to NO, to j x HCHO , and to T. Finally, a variable meant to approximate the production efficiency of ozone, $P_o(O_3)$ / $L(NO_2)$ does not follow a power-law relationship, but rather has a broad maximum at the central levels of the pollution product, (to j x HCHO x NO), at least in these Queens observations. Lower efficiency is apparent at higher and lower values of the pollution product.

The forms of the contour graphs and relationships suggest interpretation in terms of practical smog ozone control. While they may be useful, they only begin to decompose the problem. Other questions remain. How do emissions of NO_x influence NO at the measurement spot? How do organics in the atmosphere react to yield the spot VOC activity that may be approximately measured by the formaldehyde reactivity (to j x HCHO)? How do the competing processes of NO_x removal and O_3 production, and other processes, influence the total amount of ozone made, especially as air from cities mixes and moves downwind? Still, there may be a vocabulary to analyse these questions in the relationships that are shown here, and techniques complementary to idealized reacting plume analyses and massive budgets of volume-integrated transformation rates in complete urban simulation models.

Acknowledgements: We express our appreciation to Lawrence Friedl and NASA Applications *Rapid Development* program PMTACS and ARCTAS teams Laura Iraci,and internal reviewers, audience feedback at the Second International Conference on Atmospheric Chemical Mechanisms, Davis, CA, USA, 2008.

Atkinson R., and J. Arey, 2003. Gas-phase tropospheric chemistry of biogenic volatile organic compounds: a review, *Atmospheric Environment, 37* (Suppl. 2), pp. 197–219.

Atkinson, R., D. L. Baulch, R. A. Cox, J. N. Crowley, R. F. Hampson, R. G. Hynes, M. E. Jenkin, M. J. Rossi, and J. Troe, 2006. Evaluated kinetic and photochemical data for atmospheric chemistry: Volume II D. Gas phase reactions of organic species, *Atmospheric Chemistry and Physics, 6,* 3625-4055.

Brune, W.H., Faloona, I.C., Tan, D., Weinheimer, A.J.,Campos, T., Ridley, B.A., Vay, S.A., Collins, J.E., Sachse, G.W., Jaegl!e, L., Jacob, D.J., 1998. Airborne in situ OHand HO2 observations in the cloud-free troposphere and lower stratosphere during SUCCESS. *Geophysical Research Letters, 25,* 1701–1704.

Carter, W., et al., 1997. Smog chamber studies of temperature effects in photochemical smog, Environ. Science and Technol., 13, 1094–1100.

Chatfield, R.B., et al. 2009. Controls on Pollution Ozone Production Measurable from Surface, Aircraft, and Satellite Monitors, submitted for publication, *Geophys. Res. Lett.*

Eisele, F.L., Mount, G.H., Tanner, D., Jefferson, A., Shetter, R., Harder, J.W., Williams, E.J., 1997. Understanding the production and interconversion of the hydroxyl radical during the tropospheric OH photochemistry experiment. *J. Geophysical Research, 102*, 6457–6465.

Haagen-Smit, A.J., 1952. Chemistry and physiology of Los Angeles smog. *Industrial and. Engineering Chemistry, 44,* 1362.

Hastie, T., R. Tibshirani, and J. Friedman, 2001. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction,* Springer Verlag.

Jenkin, M.E., Saunders, S.M., Pilling, M.J., 1997. The tropospheric degradation of volatile organic compounds: a protocol for mechanism development, *Atmospheric Environment 31,* 81–104.

Jacobson, M., 1998. *Fundamentals of Atmospheric Modeling*, Cambridge: University Press.

Kleinman, L.I., 2005a. The dependence of tropospheric ozone production rate on ozone precursors, Atmospheric Environment 39 (2005), pp. 575–586.

Kleinman, L.I., et al, 2005b. A comparative stduy of ozone production in five U.S. metropolitan areas, J. Geophys. Res., 110, doi:10.1029/2004JD005096.

Leighton, P.A., 1961. *Photochemistry of Air Pollution*, Academic, Press, San Diego, Calif.

Lin, X., M. Trainer, S. C. Liu, On the nonlinearity of tropospheric ozone production, *J. Geophys. Res., 93,* 15,879-15,888 ,1988.

NAAMS Draft, *Draft National Ambient Air Monitory Strategy, Office of Air Quality Planning and Standards,* Research Triangle Park, NC, December, 2005.

Ren, X., Harder, H., Martinez, M., Lesher, R.L., Oliger, A.,Shirley, T., Adams, J., Simpas, J.B., Brune, W.H., 2003. HO_x concentrations and OH reactivity observations in New York City during PMTACS-NY2001, *Atmospheric Environment, 37*, 36513627–3637,\.

Ren, X., Harder, H., Martinez, M., Lesher, R.L., Oliger, A., Simpas, J.B., Brune, W.H., Schwab, J.J., Demerjian, K.L., He, Y., Zhou, X., Gao, H., OH and HO_2 chemistry in the urban atmosphere of New York City, *Atmospheric Environment, 37,* 3638-3639.

Ren, X., J. R. Olson, J. H. Crawford, W. H. Brune, J. Mao, R. B. Long, Z. Chen, G. Chen, M. A. Avery, G. W. Sachse, J. D. Barrick, G. S. Diskin, L. G. Huey, A. Fried, R. C. Cohen, B. Heikes, P. O. Wennberg, H. B. Singh, D. R. Blake, R. E. Shetter, 2003b. HO_x Chemistry during INTEX-A 2004: Observation, model comparison and comparison with previous studies, 2008. *Journal of Geophysical Research., 113*, D05310, doi:10.1029/2007JD009166.

Shafer, J. T., and J. Seinfeld, Evaluation of Chemical Reaction Mechanisms for Photochemical Smog. Part 3. Sensitivity of EKMA (Empirical Kinetic Modeling Approach) to Chemical Mechanism and Input Parameters, U.S. Environmental Protection Agency, Washington, D.C., EPA/600/3-85/042 (NTIS PB85210888, 1985).

Seinfeld, J.H., and S.N. Pandis, 1998. *Atmospheric Chemistry and Physics*, J. Wiley, Hoboken

Sillman, S., D. He, Some theoretical results concerning O_3-NO_x-VOC chemistry and NO_x-VOC indicators, *Journal of Geophysical Research, 107*, 10.1029/2001JD001123, 2002.

Singh, H. B., et al., Reactive nitrogen distribution and partitioning in the North American troposphere and lowermost stratosphere *Journal of Geophysical Research, 112*, D12S04, doi:10.1029/2006JD007664, (2007).

Thornton, J.A. et al., Ozone production rates as a function of NO_x abundances and HO_x production rates in the Nashville urban plume, 2002 *Journal of Geophysical Research, 107*, 4146, doi:10.1029/2001JD000932).W.N.

U.S. EPA, U.S. EPA, Guidelines of Use of the City-Specific EKMA in Preparing Post-1987 SIP'S , Office of Air Quality Planning and Standards, Research Triangle Park, North Carolina, 1987.

Wood. S., 2006. *Generalized Additive Models: An Introduction with R,* Chapman and Hall.

Venables, W.N. B.D. Ripley, 2004. *Modern Applied Statistics with S*, (Springer Verlag, New York,) pp 38–243.

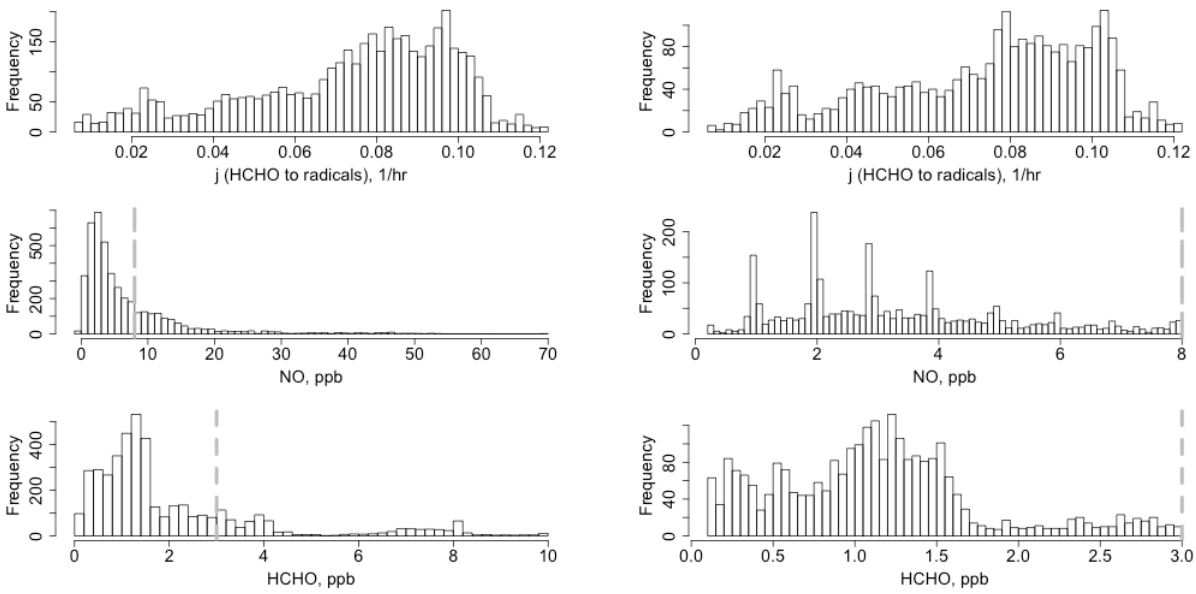
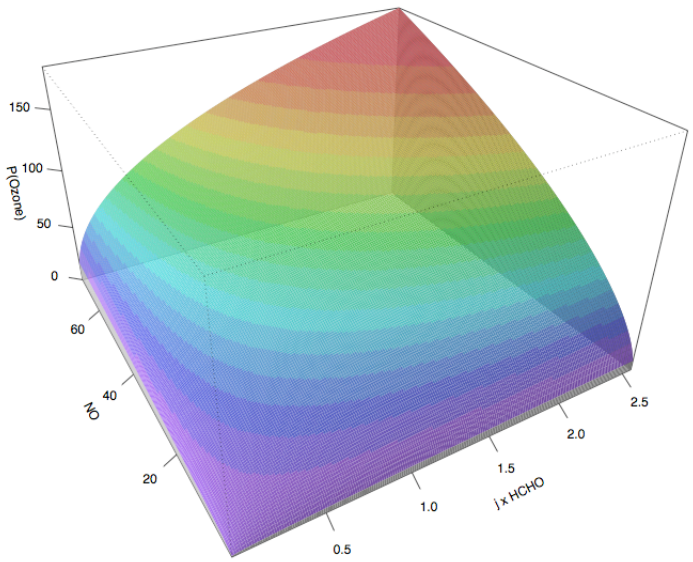
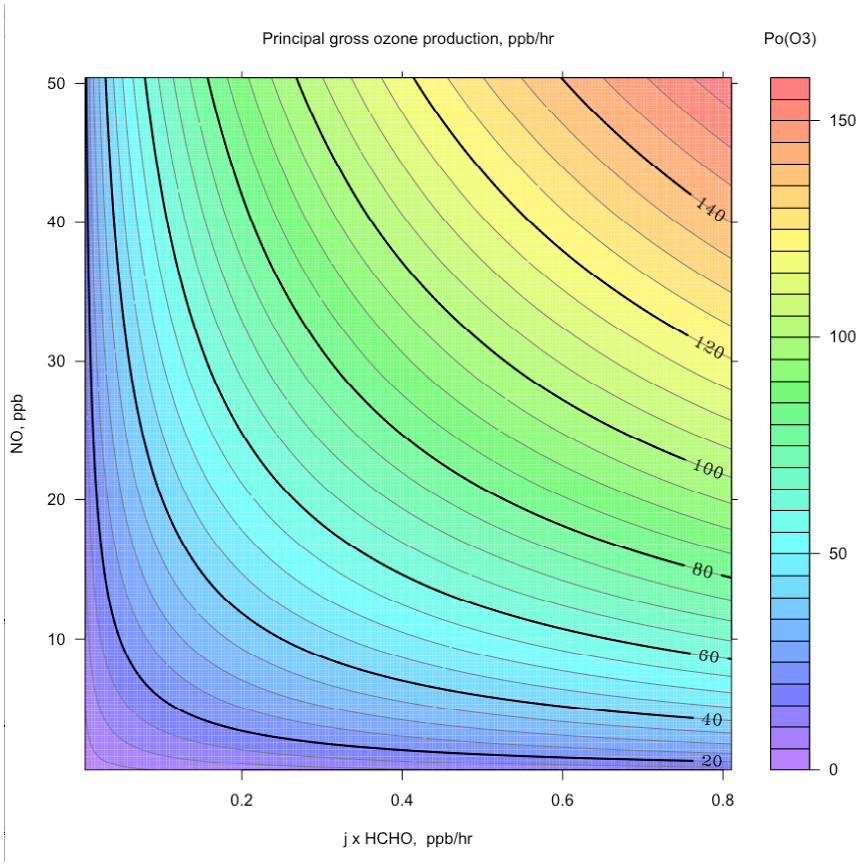


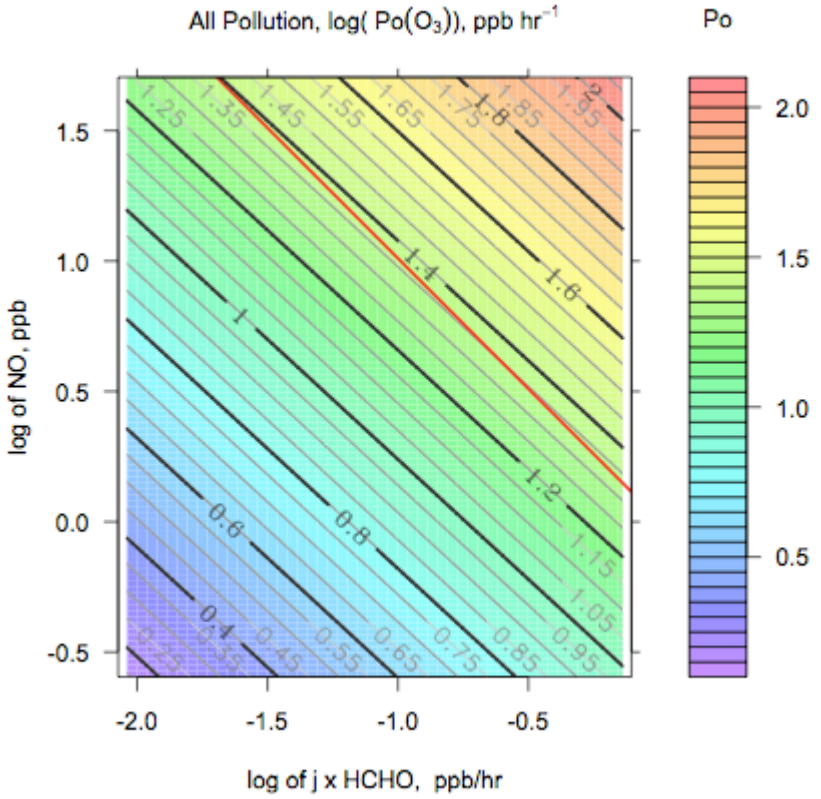
Figure 1. Histograms of observed distributions of important quantities for production of ozone as sampled at Queens College, in (a) all complete samples including highly polluted cases and (b) in a selection of "moderate pollution" cases specified by the condition that NO concentrations be less than 8 ppb (vertical long dashed line). Although this constrained HCHO also, there the ratio [NO] / [HCHO] varied widely. The NO concentrations in (b) show a preference for even-ppb values, considered an artifact of recording. This artifact may have limited achievable correlation of statistical fits with observations somewhat.


P(Ozone)
150
100
50
0
NO
60
40
20
j x HCHO
0.5
1.0
1.5
2.0
2.5

(a)

670

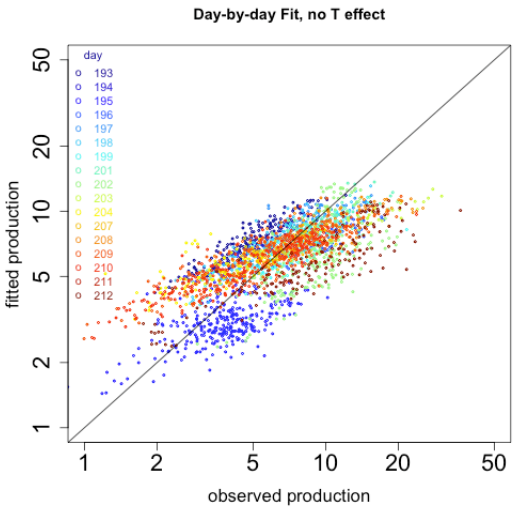

Principal gross ozone production, ppb/hr
Po(O3)
NO, ppb
50
40
30
20
10
140
120
100
80
60
40
20
150
100
50
0
0.2
0.4
0.6
0.8
j x HCHO, ppb/hr



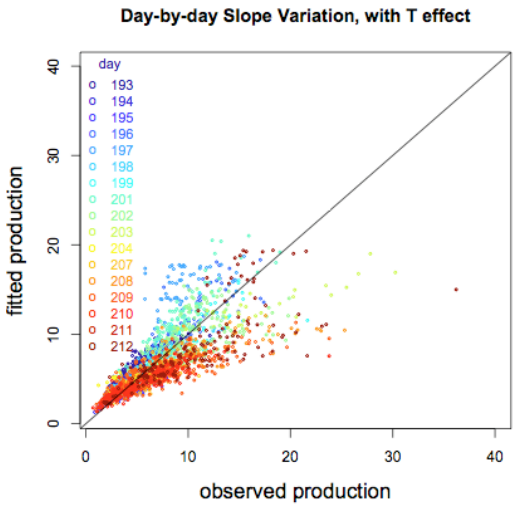
(b)

(c)

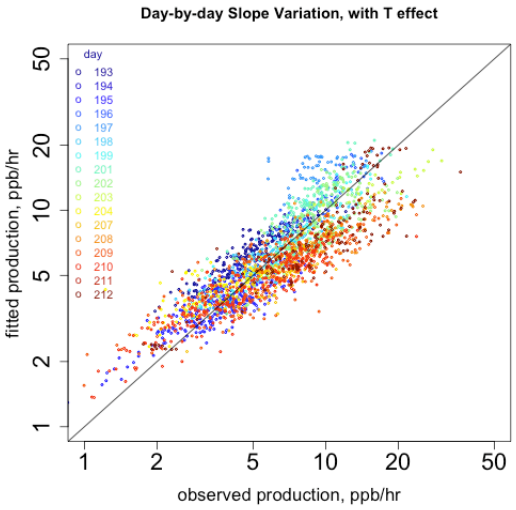
Figure 2. Ozone production as a function of organic activity as measured by HCHO photolysis and of NO, portrayed in several ways. This estimate considers all pollution samples measured at Queens College, including high-pollution cases. See text for power-law estimation procedure. (a) Contours shown on a thee-dimensional surface, emphasiing the nature of diminishing production rate as a function of each coordinates: units are in ppb and hours; i.e., $P_o(O_3)$ in ppb/hr, NO in ppb, and HCHO activity in ppb/hr. (b) The same information displayed in the contour-plot style used most commonly with Empirical Kinetic Model A (EKMA) methods, also in ppb units. (c) A plot using logarithmic quantities, with the abscissa and ordinate shown in equal aspect ratio, i.e., equal relative sensitivity $\partial \log 10 (PO_3) / \partial \log 10$ (coordinate). It is convenient to use the molecular units, molec cm-3 and s-1; all rates and statistical fits are made in these units.



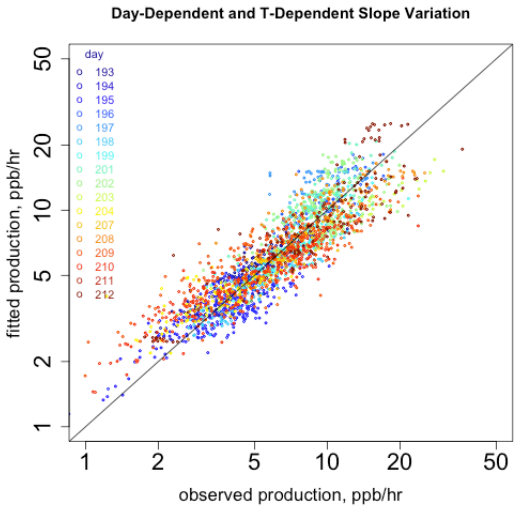
(a)



(b)



(c)



(d)

(b)

690

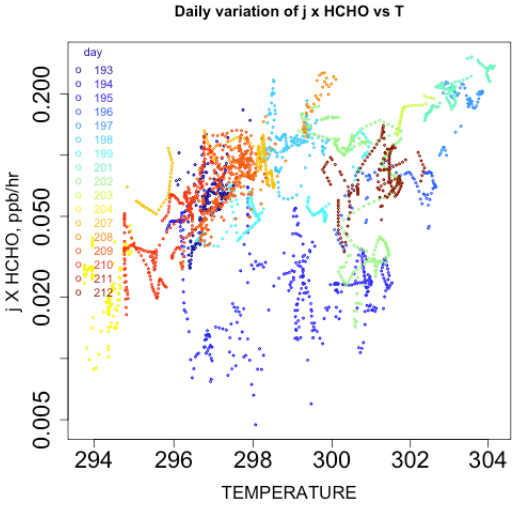


Figure 3. (a) Comparison of prediction with only j x HCHO and NO to observations based on HO_2 +NO, log scale. (b) Comparison of prediction including also *T*/298, to observations, linear scale exhibiting different slope on different days. (c) Same prediction including *T*/298, log scale. (d) Comparison of prediction including a slowly changing day-by-day factor, log scale. (e) Complex relationship of *T* and $\log_{10}$(j x HCHO).

(e)

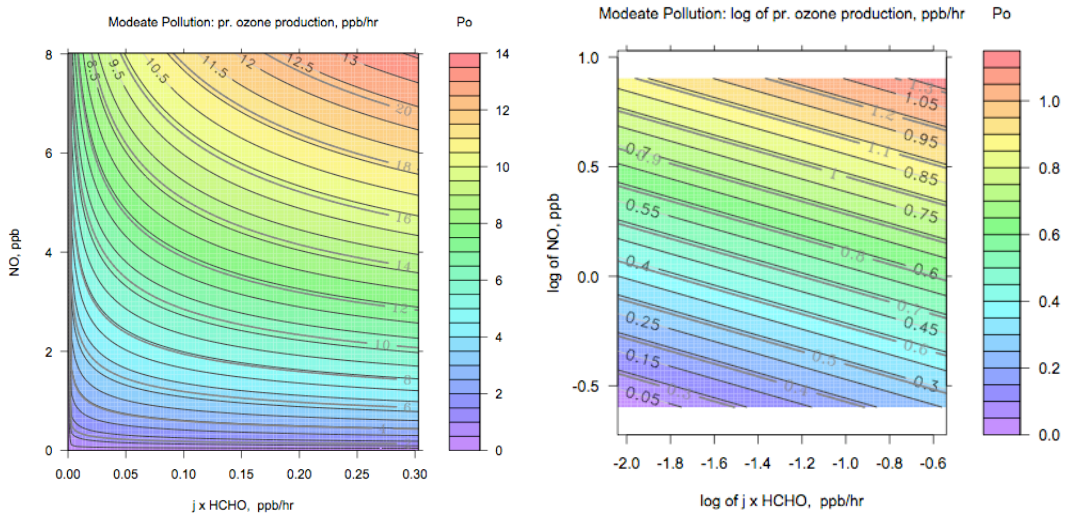


Figure 4. Ozone production as a function of organic activity as measured by HCHO photolysis and of NO, portrayed in several ways. Only moderate pollution samples are considered here (NO ≤ ??). (a) Contours shown in linear terms. (b) Contours shown in log terms; axes are set up at a 1:1 ratio. For each graph, the heavy, color-filled contours (and side-scale) indicate the fit at T = 298 K, approximately the median temperature. The gray contours that are not color-filled indicate a very much higher production rate at the 90th-quantile temperature (10% of the T measurements are higher).

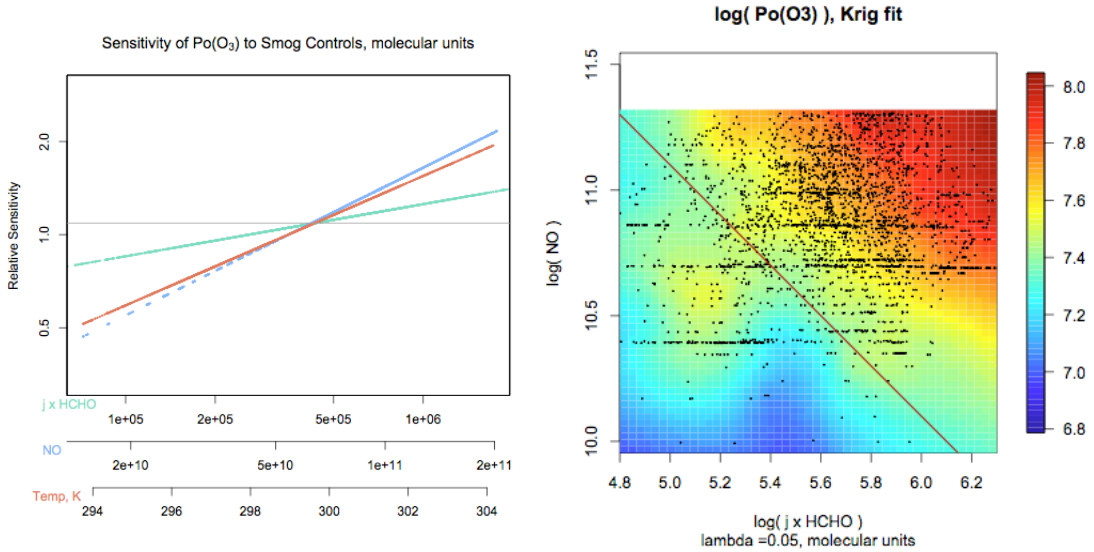


Figure 5. Response of $P_o(O_3)$ to identified controlling variables; range of variables shown includes 98% of observations. Variation with temperature at this surface site and period is roughly comparable to the variability with to j x HCHO in air-borne measurements that

we have analyzed. (b) a depiction of observed $P_o(O_3)$ in terms of j x HCHO and NO, all in log log terms and using molecular units.

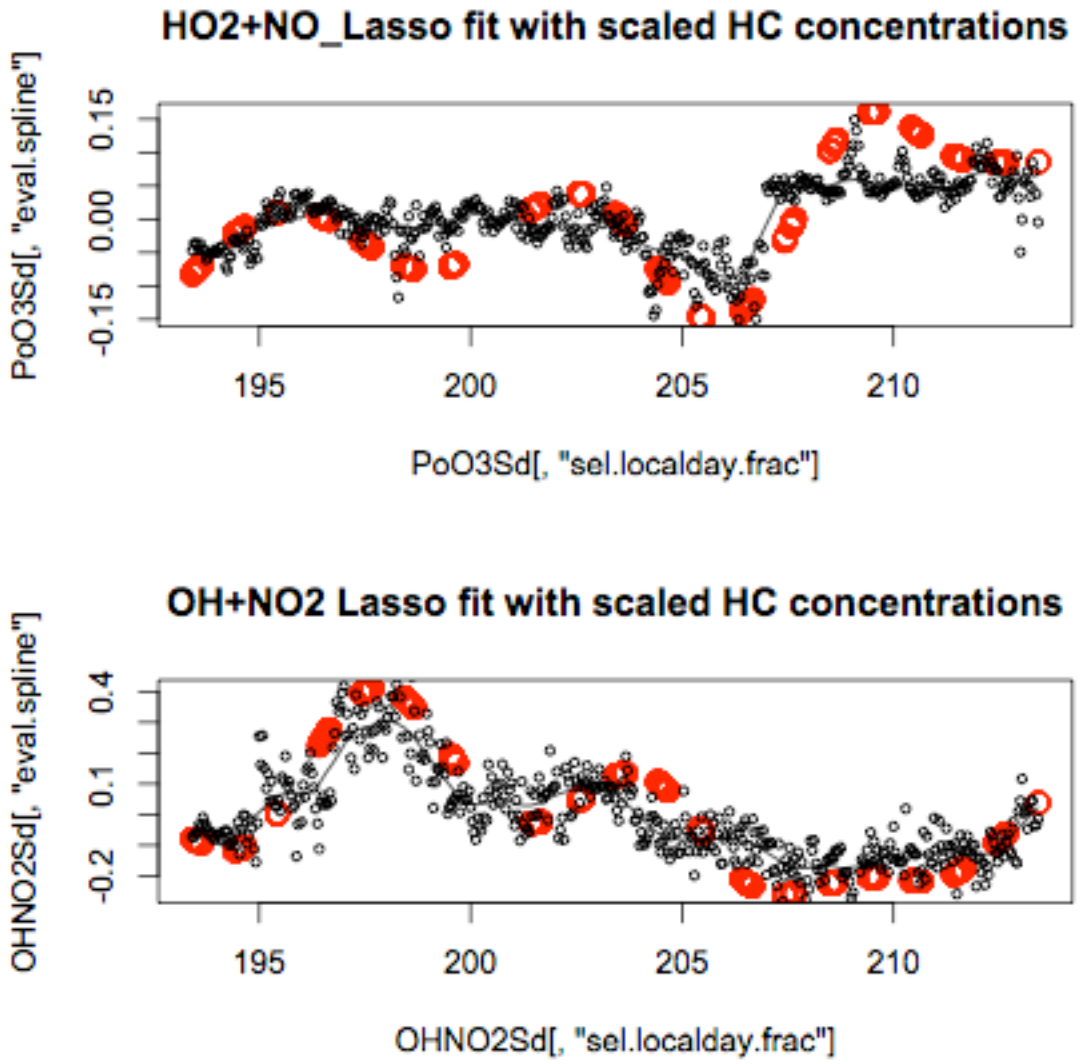


Figure 6. (a) Red circles. Slow variation of a time-dependent fitting parameter for $P_o(O_3)$ which greatly improves the correlation of statistical fit with measurements. Smaller dark circles indicate variation of a best-fitting linear combination of hydrocarbon measurements chosen to mimic the trend of the red circles. (b) A similar depiction of a fit of the loss rate of NO_2 due to OH. Red circles are fitted smooth t-variation (see) text, dark circles are best-fit linear combination of hydrocarbon concentrations. These hydrocarbons may indicate a sensitivity to VOC reaction-rate with OH.

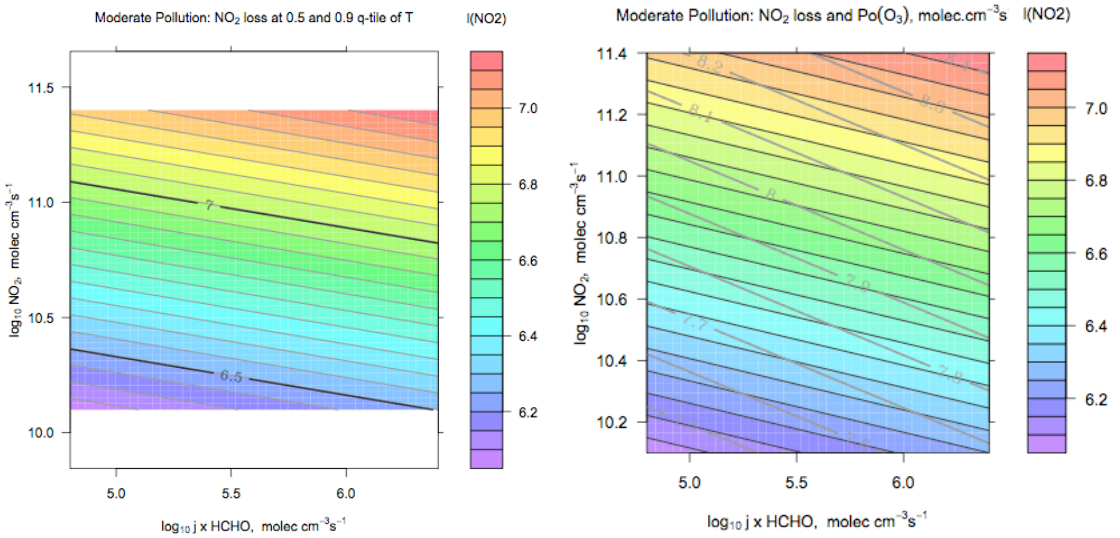


Figure 7. Loss of nitrogen oxides by OH+NO_2 reaction, shown in color-filled contours. (a) Graph showing a large sensitivity to NO_2 and a moderate sensitivity to T. The heavy lines show the dependence at 301.5 K, about five contour lines displaced from the color-graph scale, corresponding to T =298 K. (b) Graph does not have 1:1 aspect ratio. The gray contours indicate for comparison the sensitivity of the production rate of ozone; to do this, NO was estimated from NO_2 using a simple rescaling of the ratio of means. L(NO_2) is very NO_x-sensitive. This graph does not represent all the complexities of the relationships of $P_o(O_3)$ and L(NO_2).

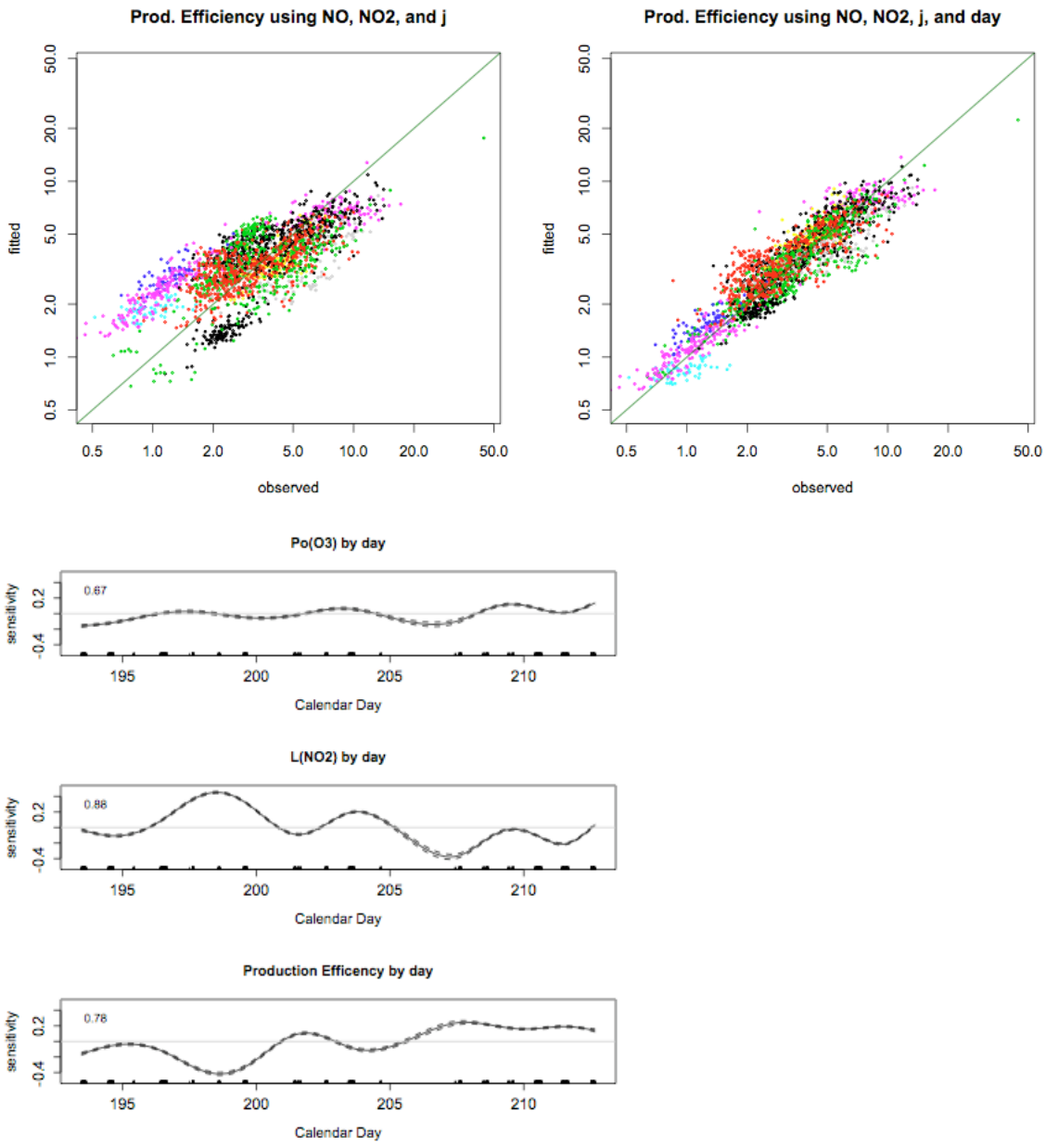


Figure 8. (a) Calculated and fitted production officiency of ozone (dimensionless) as described by $Po(O_3)$, $L(NO_2)$ using variables useful for prediction. Note the day-to-day variation shown by different colors (b) Similar comparison using also an empirical fit that describes a slow variation through the measurement. (c) Slow, multiday variation of sensitivity function *s* for $Po(O_3)$, $L(NO_2)$, and a direct empirical estimate of an approximate "ozone production efficiency," $Po(O3) / L(NO_2)$. The production efficiency variation mirrors the variation for photochemical NO_x removal, $OH+NO_2$.